(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 847 003 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
10.06.1998 Bulletin 1998/24

(51) Int Cl.[6]: **G06F 3/16**

(21) Application number: **97309707.4**

(22) Date of filing: **02.12.1997**

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **03.12.1996 US 32507 P**

(71) Applicant: **TEXAS INSTRUMENTS INC.**
Dallas, Texas 75243 (US)

(72) Inventors:
- **Oh, Stephen S.**
  Richardson, Texas 75082 (US)
- **Popik, Stephen Ira**
  Plano, Texas 75093 (US)

(74) Representative:
Legg, Cyrus James Grahame et al
ABEL & IMRAY,
Northumberland House,
303-306 High Holborn
London WC1V 7LH (GB)

(54) An audio memo system and method of operation thereof

(57) The hands-free audio memo system (10) and method uses an interface (12) to receive a voice input from a user, and a speech recognition unit (18) coupled to the interface (12) to monitor the voice input and recognize a predetermined set of voice commands from the voice input. The speech recognition unit (18) generates a command signal that corresponds to the recognized voice command, which is received by a controller unit (20). The controller unit (20) activates a speech acquisition unit (16) coupled to the controller unit (20) to collect and stop collecting the voice input in response to a control signal generated by the controller unit (20). a memory (24) is provided to store the collected voice input.

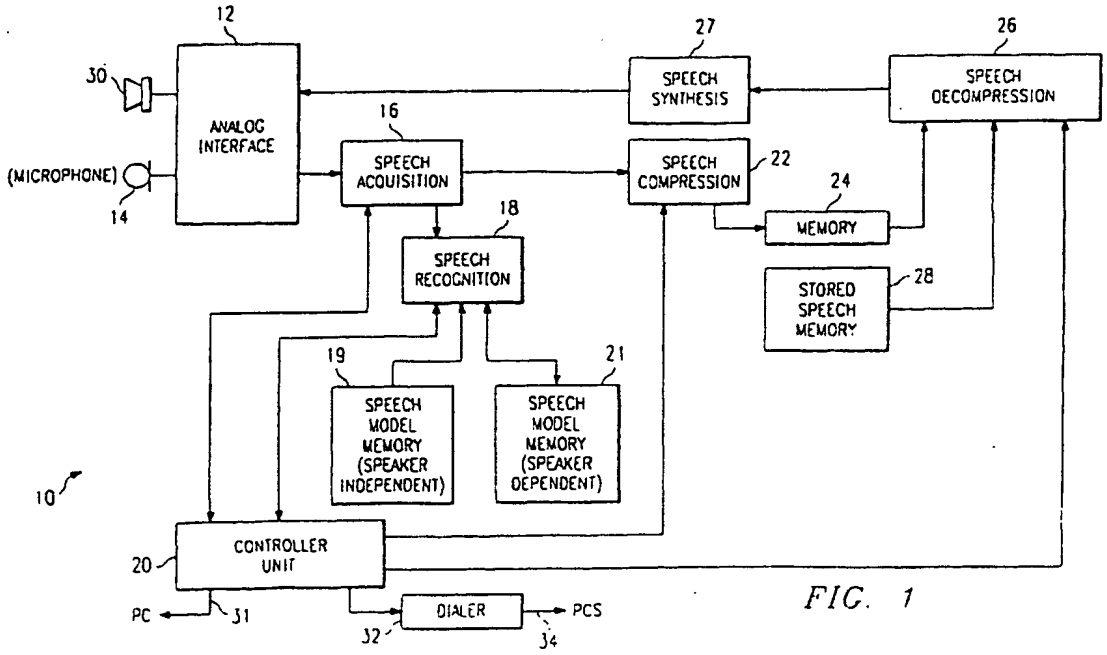


FIG. 1

EP 0 847 003 A2

Description

TECHNICAL FIELD OF THE INVENTION

This invention is related in general to the field of personal electronic systems. More particularly, the invention is related to an audio memo system and method of operation thereof.

BACKGROUND OF THE INVENTION

It is common knowledge that we are currently living in the Information Age. Data comes to us in visual, audio, and written forms through a myriad of channels: radio, telecommunications, television, internet, world wide web, and just plain seeing, hearing, and feeling things as events occur around us. There are many instances when it is desirable to retain some of the information in a more reliable manner than the ability or inability to recall data we are born with. For example, a telephone number announced on the radio, the location of a specialty store, or an ingenious idea about a novel gadget to solve a stubborn problem.

The old standby to record data is the pen and paper. However, there are times when it is inconvenient to write, such as when one is operating an automobile, or when pen and paper are not accessible.

Dictaphones, which use audio tape cassettes, and some newer digital recorders, have been used to fill this void. However, they all require the use of at least one hand to hold the device, and to operate one of many buttons to turn "ON" the device, record, retrieve, erase, and turn "OFF" the device. Further, because it has been shown that the use of one hand to handle a wireless telephone while operating an automobile can lead to unsafe driving and possibly higher incidents of traffic accidents, it is less than desirable to also require the driver to devote the use of one hand to operate the recording device.

SUMMARY OF THE INVENTION

Accordingly, there is a need for an audio memo system which enables hands-free operation.

In accordance with the present invention, a hands-free audio memo system and method of operation thereof are provided which eliminate or substantially reduce the disadvantages associated with prior devices.

In one aspect of the invention, an audio memo system is provided that uses an interface for receiving a voice input from a user, and a speech recognition unit coupled to the interface for monitoring the voice input and recognizing a predetermined set of voice commands from the voice input. The speech recognition unit generates a command signal that corresponds to the recognized voice command, which is received by a controller unit. The controller unit activates a speech acquisition unit coupled to the controller unit for collecting and or stopping collection of the voice input in response to a control signal generated by the controller unit. A memory is provided for storing the collected voice input.

In another aspect of the invention, the personal memo system includes an analog interface for receiving a voice input from a user, a speech recognition unit coupled to the interface and adapted for recognizing a predetermined set of voice commands from the voice input, and for generating a command signal in response thereto. A controller unit is coupled to the speech recognition unit which generates a control signal in response to receiving the command signal from the speech recognition unit. A digital telephone answering device is coupled to the controller unit and analog interface for collecting and storing the voice input.

In yet another aspect of the invention, a method for operating an audio memo system includes the steps of receiving a voice input from a user, recognizing voice commands in the voice input indicative of the user's desire to record an audio memo, collecting subsequent voice input, and storing the subsequent voice input.

The audio memo system of the present invention provides a way for users to record audio memos and perform other functions without the use of a hand for its operation. This is especially advantageous for persons who are operating an automobile or performing other tasks that require concentration and generally the use of both hands.

BRIEF DESCRIPTION OF THE DRAWINGS

For a better understanding of the present invention, reference may be made to the accompanying drawings, in which:

FIGURE 1 is a simplified functional block diagram of an exemplary hands-free audio memo system constructed according to the teachings of the present invention;

FIGURE 2 is a simplified block diagram of an alternative embodiment of the hands-free audio memo system of the present invention;

FIGURE 3 is an exemplary flowchart of a simplified hands-free audio memo algorithm according to the teachings of the present invention;

FIGURE 4 is an exemplary flowchart of an hands-free audio memo algorithm according to the teachings of the present invention;

FIGURE 5 is an exemplary flowchart of voice playback and memo management functions of the hands-free audio memo algorithm according to the teachings of the present invention; and

FIGURE 6 is an exemplary flowchart showing exemplary voice inputs to the system according to the teachings of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

A preferred embodiment of the present invention is illustrated in FIGURES 1-6, like reference numerals being used to refer to like and corresponding parts of the various drawings.

Referring to FIGURE 1, a functional block diagram of an exemplary hands-free audio memo system 10 constructed according to the teachings of the present invention is shown. System 10 includes an analog interface 12 which receives voice input of a user captured by a microphone 14, and converts the analog voice input into a digital voice input signal. Analog interface 12 is further coupled to a speech acquisition unit 16, which functions to collect the digital voice input signal. The collected digital voice input signal is then provided to a speech recognition unit 18, which receives the digital voice input signal and searches for a set of predetermined voice commands and responses stored in a speaker-independent speech model memory 19 and/or an optional speaker-dependent speech model memory 21. For example, the voice command may be "MEMO START" or "TAKE MEMO" to initiate memo recording, "MEMO TERMINATE" to stop memo recording, and other appropriate responses. Further, certain commands and responses may only be valid during certain times and ignored at other times. For example, when memo recording is taking place, speech recognition unit 16 may only listen for a smaller set of commands and/or responses from the user, such as "MEMO TERMINATE," and not "YES" or "NO."

Speech recognition unit 18 is further coupled to a controller unit or microcontroller unit (MCU) 20. When speech recognition unit 16 recognizes a valid command or response, it generates a signal to inform controller unit 20 to take appropriate actions. Controller unit 20 is further coupled to a speech compression unit 22, which is also coupled to speech acquisition unit 16. Speech compression unit 22 compresses the digital voice input signals collected by speech acquisition unit 16 using known compression algorithms and stores the compressed signals into a memory 24.

A speech decompression unit 26 and a speech synthesis unit 27 are further coupled between controller unit 20 and analog interface 12. Controller unit 20 instructs speech compression unit 26 to decompress stored speech in memory 28 and provide to speech synthesis unit 27 to produce a speech prompt or response at appropriate times, which is then broadcast to the user by a speaker 30 coupled to analog interface 12.

Optionally, a communications link 31 may be provided to download voice input signals stored in memory 24 to a personal computer (not shown). In addition, a dialer 32 and link 34 may be further provided to a personal communications system (not shown) to perform functions related to telecommunications, such as dialing a particular number or "CALL HOME."

FIGURE 2 is a simplified block diagram of an embodiment of hands-free audio memo system 50 according to the teachings of the present invention. System 50 includes an analog interface 52 coupled to a speech recognition unit 54 and a digital telephone answering device (DTAD) 56. DTAD 56 typically includes speech acquisition and compression functions, and a memory. A microcontroller unit 58 is further coupled to speech recognition unit 54 and DTAD 56.

System 50 may be implemented with commercially available components or devices. For example, interface 52 may be implemented with TCM320AC36 or TCM320AC37 Voice-Band Audio Processors (VBAP)™ manufactured by Texas Instruments Incorporated of Dallas, Texas; speech recognition unit 54 may be implemented with TMS320C5X Digital Signal Processor (DSP) also manufactured by Texas Instruments Incorporated; DTAD 56 may be implemented with the MSP58C8X product line of Texas Instruments Incorporated; and microcontroller unit 58 may be implemented with TMS370 family products of Texas Instruments Incorporated.

A single chip implementation of the audio memo system is also contemplated. For example, components in Texas Instrument's cDSP™ product line may be incorporated and formed on a single silicon substrate to construct an integrated circuit. For example, a C54X core for performing the speech recognition and DTAD functions, an Advanced RISC (reduced instruction set computing) Machines (ARM™) 7TDMI core for performing the controller unit functions, and a Voice-Band Audio Processor core for performing analog interface functions may be combined into a single integrated circuit. It may be seen that the above are merely examples and other suitable substitutes may be used.

Referring to FIGURE 3 as well as the block diagrams in FIGURES 1 and 2, an exemplary process flow 70 for hands-free audio memo systems 10 and 50 is provided. Speech acquisition 16 or DTAD 56 and recognition 16 or 54 is first activated in step 72. For example the activation may be done at the time the automobile (not shown) is started, by the push of a button, or by leaving the key in the accessory position, for example. In steps 72 and 74, speech recognition unit 18 or 54 searches for a valid command appropriate for the occasion, such as "MEMO START" to start the memo recording process. Once a valid command is recognized, as determined in step 76, controller unit 20 or 58 is notified, such as by a signal generated by speech recognition unit 18 or 54, as shown in step 78. Controller unit 20 or 58 then activates the memo function, as shown in step 80. Once the system is ready, an optional audio prompt or speech (e.g., "MEMO SYSTEM READY") may be generated to signal to the user that he/she may begin to speak. A timer or counter (not shown) set for a predetermined time period may be started when speech acquisition 16 begins to capture voice input. The collected voice input is converted to digital signals, compressed and stored in memory 24, as shown in step 86. When

the timer expires, speech acquisition is stopped, as shown in step 88. Controller unit 20 or 58 is then notified that memo recording terminated, as shown in step 90, and execution returns to step 74 to be ready for the next memo.

A further, alternative method for hands-free audio memo 100 is shown in FIGURE 4. Speech acquisition 16 or DTAD 56 and speech recognition 18 or 54 are activated either by starting the automobile, leaving the key in the accessory position, or the push of a button (not shown), for example, as shown in step 102. Speech recognition 18 or 54 monitors the speech uttered by the user(s) in the vicinity and searches for recognizable valid voice commands and/or responses, such as a command to start the memo process, as shown in step 104.

When it is determined that the captured voice input is a valid command, such as "MEMO START," controller unit 20 or 58 is notified, as shown in steps 106 and 108. Controller unit 20 or 58 in turn activates the memo function, as shown in step 110. In step 112, an audio prompt or speech (e.g., "MEMO SYSTEM READY") may be generated to signal to the user that he/she may begin to speak. The user's speech is then captured and compared with recognizable commands appropriate for the circumstances, such as "MEMO TERMINATE" to end the process, as shown in steps 114 and 116. Speech recognition 18 or 54 may be running in a low resource mode at this time to look for only those commands that are valid during this time, such as only the command to terminate or pause the memo taking process. If the captured utterance is not a recognizable and valid command, then it is collected, compressed, and stored, as shown in step 118. If in step 116, it is determined that the captured speech is a recognizable and valid command to end the memo process, for example, then controller unit 20 or 58 is notified, as shown in block 120. Controller unit 20 or 58 then pauses speech acquisition, as shown in step 122, and instructs speech decompression 26 and speech synthesis 27 to issue an audible prompt for confirmation, such as "READY TO TERMINATE MEMO?" The subsequent voice input is then captured and monitored for a valid response to the prompt, such as "YES" or "NO," as shown in steps 126 and 128. If the received voice input is not a recognizable valid response to the confirmation, then an appropriate audio response may be generated to reconfirm, as shown in step 132. If the voice input is recognized as a response indicative that the user is not ready to terminate the memo process, then execution returns to step 112, to continue to record memo. If on the other hand the voice input is recognized as an affirmative response in step 130, then the memo function is stopped in step 134, and controller unit 20 or 58 is notified in step 136. Execution then returns to step 104 to prepare for the next memo.

FIGURE 5 is a flowchart of memo playback and memo management functions of system 10 and 50. At step 76 shown in FIGURE 3 or step 106 shown in FIGURE 4, if the voice input is not a valid start command, it is also checked for whether it is a valid playback command, as shown in step 140. If it is, controller unit 58 is notified in step 142 and the user is prompted for additional input, which is captured, as shown in step 144. The captured speech input is then examined to determined whether it is a valid response to the prompt given in step 144, if not, some appropriate action is taken in step 148, such as issue an appropriate audio statement. If it is a valid response, then the memo playback function 150 is launched, where the user may play back one or more previously recorded memos, skip one or more memos, etc. At the end of the memo playback function, the algorithm may retum to step 74 in FIGURE 3 or step 114 shown in FIGURE 4.

If in step 140 it is determined that the speech input is not a valid playback command, then a determination is made as to whether it is a valid memo management command in step 152. If not, then the process may return to step 74 in FIGURE 3 or step 114 shown in FIGURE 4 to continue to capture the speech input. Otherwise, controller unit 58 is notified in step 154 and the user is prompted for additional input, which is captured, as shown in step 156. The captured speech input is then examined to determined whether it is a valid response to the prompt given in step 158, if not, some appropriate action is taken in step 148, such as issue an audio statement. If it is a valid response, then the memo management function 160 is launched, where the user may perform operations such as delete, save, and protect on previously recorded memos. At the end of the memo management function, the algorithm may return to step 74 in FIGURE 3 or step 114 shown in FIGURE 4.

Referring to FIGURE 6, a more detailed process flow is shown. As voice input is captured in step 170, it is determined whether it matched any recognizable and valid command and response in step 172. For example, one or more recognized key phrases may be used to initiate system 50 in a memo recording mode 180, memo playback mode 182, memo management mode 184, dialer mode 186, and voice mail mode 188, where each mode is shown with exemplary valid phrases recognized when system 50 is in the respective modes. The key phrases to launch each mode may include "MEMO START" to launch the memo recording functions; "MEMO PLAYBACK" to launch the memo playback functions; "MEMO MANAGEMENT" to launch the memo management functions; "CALL X" to launch the dialer functions; and "GET MAIL" to launch the voice mail functions. Thus, speech recognition unit 54 need only to focus on a subset of possible valid utterances as to speed up search and processing time and to conserve resources.

Although the present embodiment and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made therein without departing from the spirit and scope of the invention.

Claims

1. An audio memo system, comprising:

   an interface for receiving a voice input from a user;
   a speech recognition unit coupled to the interface and adapted for recognizing a predetermined set of voice commands from the voice input, the speech recognition unit further being arranged for generating a command signal associated with the recognized voice command;
   a controller unit coupled to the speech recognition unit for generating a control signal in response to receiving the command signal from the speech recognition unit;
   a speech acquisition unit coupled to the controller unit and the interface for collecting or terminating collection of the voice input in response to the control signal; and
   a memory unit coupled to the speech acquisition unit for storing the collected voice input.

2. The system as set forth in Claim 1, further comprising;

   a speech compression unit coupled to the speech acquisition unit for compressing the collected voice input prior to storing in the memory unit.

3. The system, as set forth in Claim 1 or Claim 2, further comprising;

   a speech synthesis unit coupled to the interface and speech recognition unit for generating a predetermined set of audio confirmations in response to the voice input.

4. The system as set forth in any of Claims 1 to 3, wherein the controller unit comprises a communications link coupled to a personal computer for downloading stored voice inputs from the memory unit thereto.

5. The system as set forth in any of Claims 1 to 4, wherein the controller unit comprises a communications link coupled to a personal communication system for dialing a telephone number in response to a dialing command recognized by the speech recognition unit.

6. A method for operating an audio memo system, comprising the steps of:

   receiving a voice input from a user;
   recognizing a voice command in the voice input indicative of the user's desire to record an audio memo;
   collecting subsequent voice inputs; and
   storing the subsequent voice inputs.

7. The method as set forth in Claim 6 further comprising the step of;

   generating an audio confirmation in response to recognizing the voice command.

8. The method as set forth in Claim 7, wherein the audio confirmation generating step further comprises the step of synthesizing a speech signal.

9. The method, as set forth in any of Claims 6 to 8, further comprising the steps of:

   recognizing a voice command in the subsequent voice inputs that is indicative of the user's desire to stop recording the audio memo; and
   stopping collection of subsequent voice inputs.

10. The method, as set forth in Claim 9, further comprising the step of:

    prior to said step of stopping collection of subsequent voice inputs generating an audio confirmation in response to recognizing the voice command; and
    recognizing a confirmation in the voice input.

11. The method, as set forth in any of Claims 6 to 8, further comprising the steps of:

    recognizing a voice command in the subsequent voice inputs that is indicative of the user's desire to stop recording the audio memo;
    generating an audio confirmation in response to recognizing the voice command;
    recognizing a denial in the voice input; and
    continuing collection of subsequent voice inputs.

12. The method as set forth in any of Claims 6 to 11, further comprising the steps of:

    starting a timer for a predetermined time period after the voice command recognizing step; and
    stopping collection of subsequent voice inputs when the timer expires.

13. The method as set forth in any of Claims 6 to 8, further comprising the steps of:

    converting the voice input to a digital voice input; and
    compressing the digital voice input prior to storing thereof.

14. The method, as set forth in any of Claims 6 to 8, further comprising the steps of:

    recognizing a voice command in the voice input

indicative of the user's desire to playback stored voice input; and
playing back stored voice input in response to further voice input.

15. The method as set forth in any of Claims 6 to 8, further comprising:

recognizing a voice command in the voice input indicative of the user's desire to manage stored voice input; and
performing management functions in response to further voice inputs.

16. The method, as set forth in any of Claims 6 to 8, further comprising the steps of:

recognizing a voice command in the voice input indicative of a specific function; and
focusing on a subset of subsequent valid voice commands in response thereto.

17. A personal memo system comprising:

an analog interface for receiving a voice input from a user;
a speech recognition unit coupled to the interface and adapted for recognizing a predetermined set of voice commands from the voice input, the speech recognition unit further being arranged for generating a command signal associated with the recognized voice command;
a controller unit coupled to the speech recognition unit for generating a control signal in response to receiving the command signal from the speech recognition unit; and
a digital answering device coupled to the controller unit and analog interface for collecting and storing the voice input.

18. The system as set forth in Claim 17, wherein the controller unit comprises a communications link coupled to a personal computer for downloading stored voice inputs from the memory unit thereto.

19. The system as set forth in Claim 17, wherein the controller unit comprises a communications link coupled to a personal communication system for dialing a telephone number in response to a dialing command recognized by the speech recognition unit.

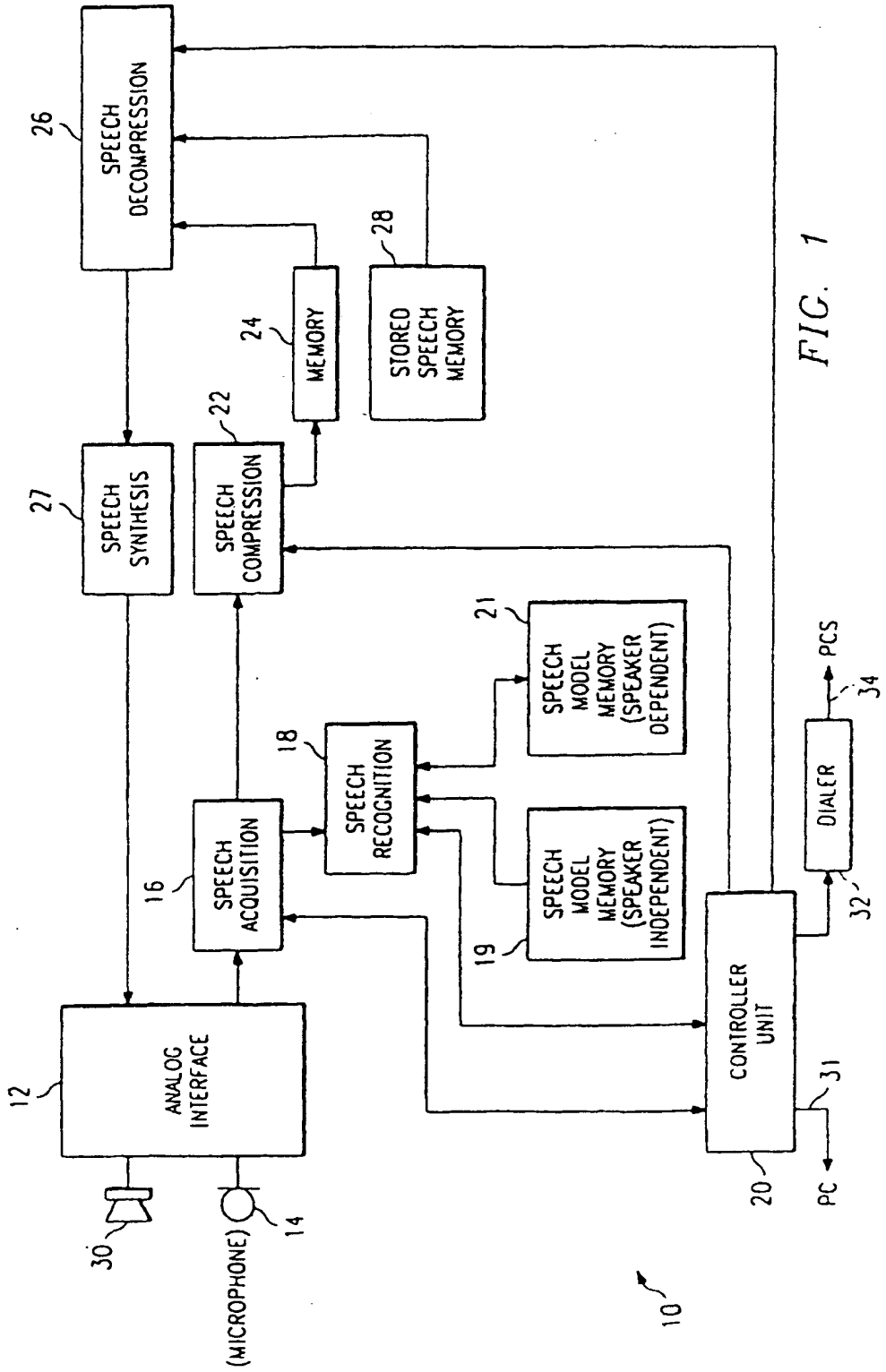

SPEECH DECOMPRESSION
26
MEMORY
24
STORED SPEECH MEMORY
28
SPEECH SYNTHESIS
27
SPEECH COMPRESSION
22
SPEECH MODEL MEMORY (SPEAKER DEPENDENT)
21
SPEECH RECOGNITION
18
SPEECH ACQUISITION
16
SPEECH MODEL MEMORY (SPEAKER INDEPENDENT)
19
DIALER
32
PCS
34
ANALOG INTERFACE
12
CONTROLLER UNIT
20
31
PC
30
(MICROPHONE)
14
10

FIG. 1

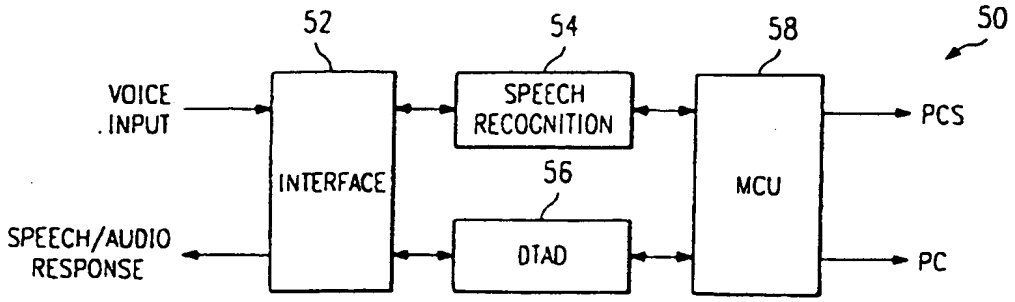

52
54
58
50
VOICE INPUT
INTERFACE
SPEECH RECOGNITION
MCU
PCS
56
SPEECH/AUDIO RESPONSE
DTAD
PC

FIG. 2

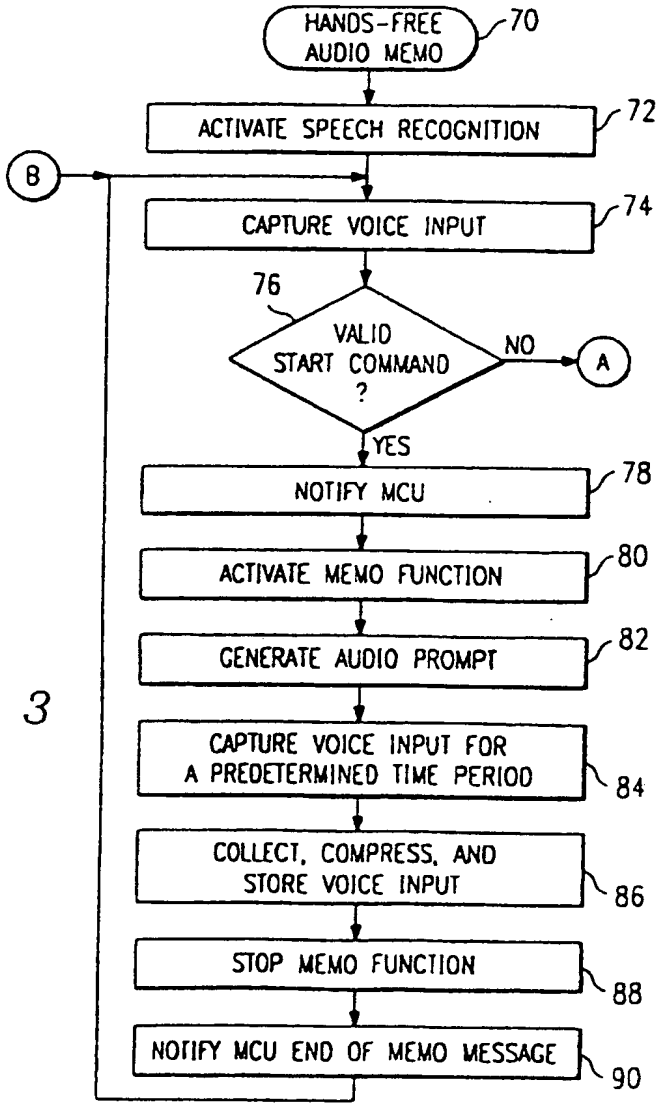
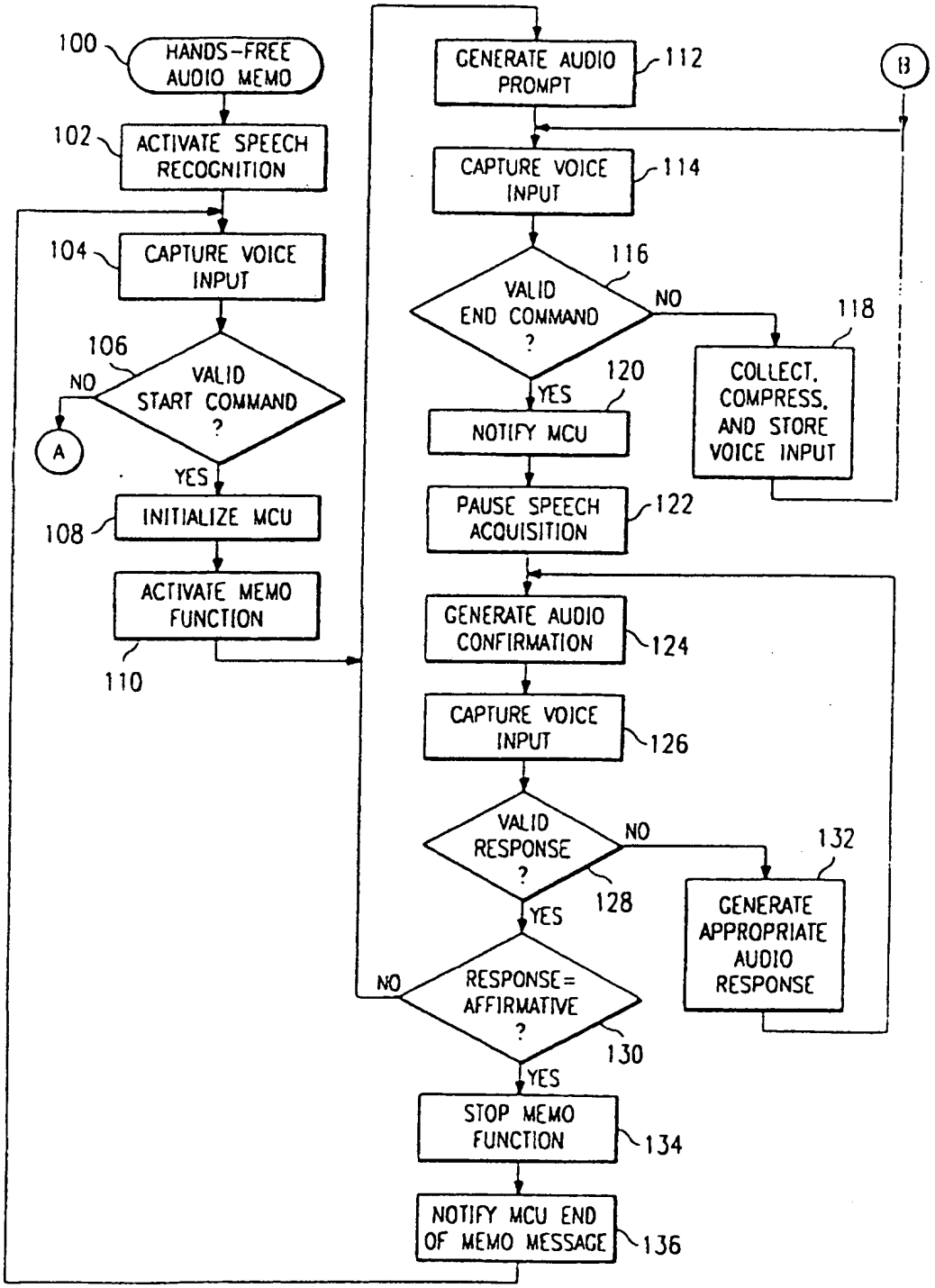

HANDS-FREE AUDIO MEMO
70
ACTIVATE SPEECH RECOGNITION
72
B
CAPTURE VOICE INPUT
74
76
VALID START COMMAND ?
NO
A
YES
NOTIFY MCU
78
ACTIVATE MEMO FUNCTION
80
GENERATE AUDIO PROMPT
82
CAPTURE VOICE INPUT FOR A PREDETERMINED TIME PERIOD
84
COLLECT, COMPRESS, AND STORE VOICE INPUT
86
STOP MEMO FUNCTION
88
NOTIFY MCU END OF MEMO MESSAGE
90

FIG. 3

FIG. 4


100
HANDS-FREE AUDIO MEMO
102
ACTIVATE SPEECH RECOGNITION
104
CAPTURE VOICE INPUT
106
VALID START COMMAND ?
NO
YES
A
108
INITIALIZE MCU
ACTIVATE MEMO FUNCTION
110
GENERATE AUDIO PROMPT
112
CAPTURE VOICE INPUT
114
116
VALID END COMMAND ?
NO
118
COLLECT, COMPRESS, AND STORE VOICE INPUT
YES
120
NOTIFY MCU
PAUSE SPEECH ACQUISITION
122
GENERATE AUDIO CONFIRMATION
124
CAPTURE VOICE INPUT
126
VALID RESPONSE ?
NO
132
GENERATE APPROPRIATE AUDIO RESPONSE
128
YES
NO
RESPONSE = AFFIRMATIVE ?
130
YES
STOP MEMO FUNCTION
134
NOTIFY MCU END OF MEMO MESSAGE
136
B

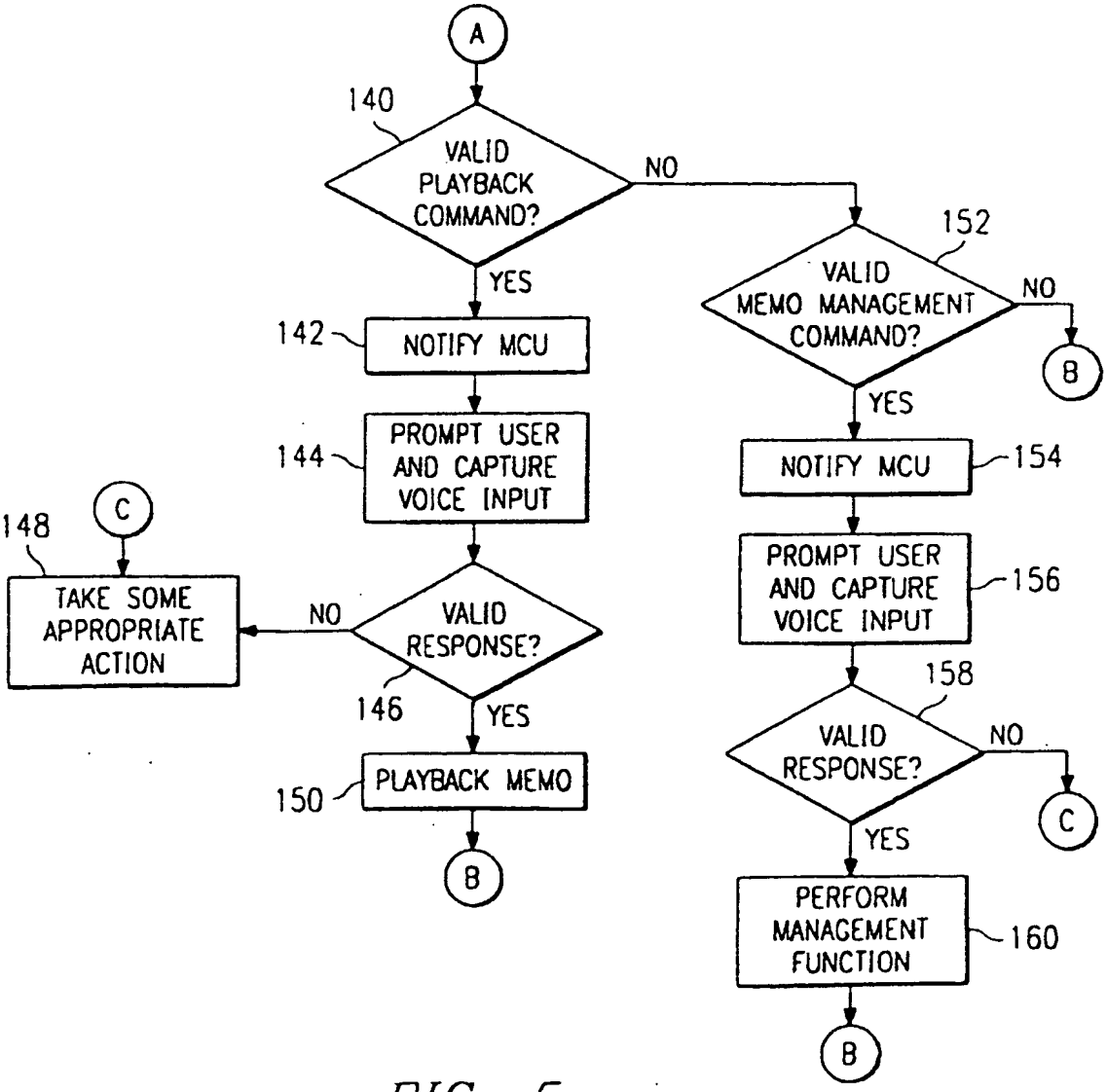

A
140
VALID PLAYBACK COMMAND?
NO
YES
142
NOTIFY MCU
144
PROMPT USER AND CAPTURE VOICE INPUT
148
C
TAKE SOME APPROPRIATE ACTION
NO
VALID RESPONSE?
146
YES
150
PLAYBACK MEMO
B
152
VALID MEMO MANAGEMENT COMMAND?
NO
B
YES
NOTIFY MCU
154
PROMPT USER AND CAPTURE VOICE INPUT
156
158
VALID RESPONSE?
NO
C
YES
PERFORM MANAGEMENT FUNCTION
160
B

FIG. 5

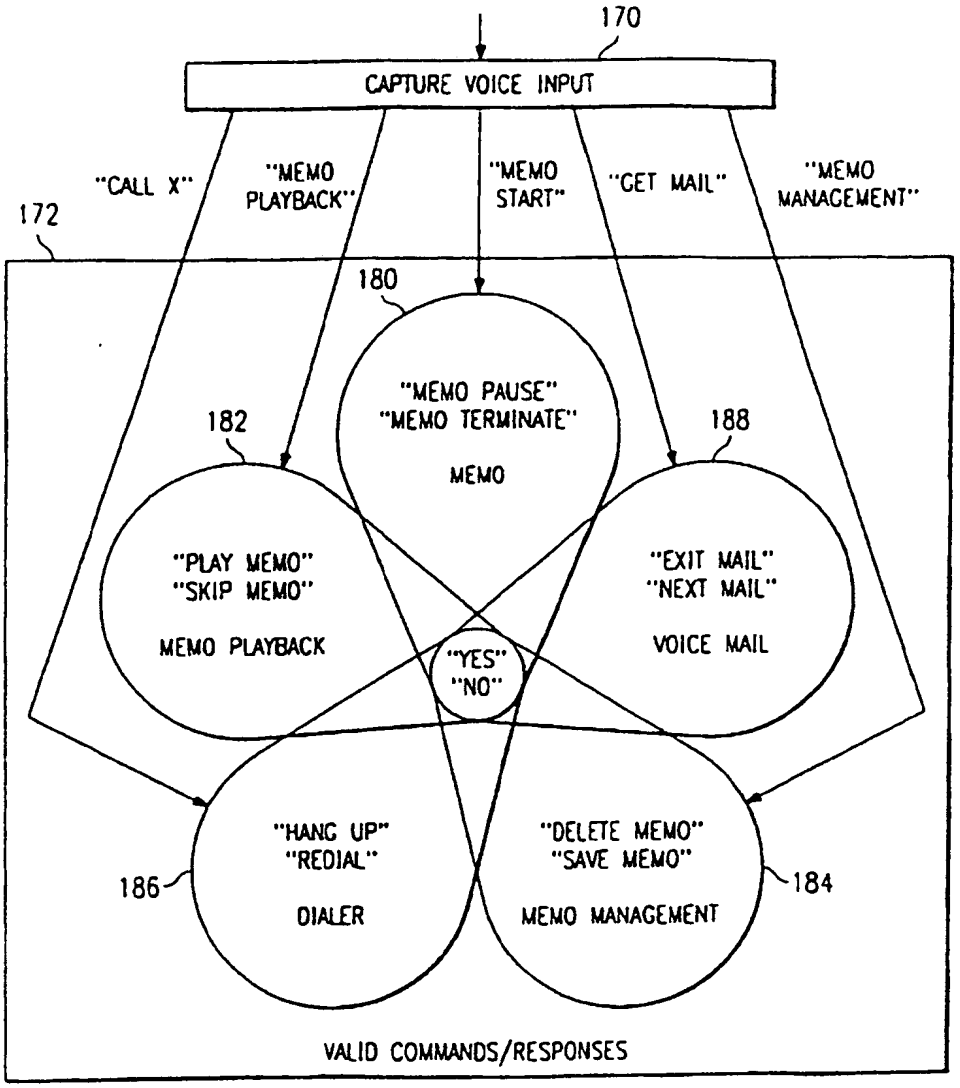

170
CAPTURE VOICE INPUT
"CALL X"
"MEMO PLAYBACK"
"MEMO START"
"GET MAIL"
"MEMO MANAGEMENT"
172
180
"MEMO PAUSE"
"MEMO TERMINATE"
MEMO
182
188
"PLAY MEMO"
"SKIP MEMO"
MEMO PLAYBACK
"YES"
"NO"
"EXIT MAIL"
"NEXT MAIL"
VOICE MAIL
"HANG UP"
"REDIAL"
DIALER
"DELETE MEMO"
"SAVE MEMO"
MEMO MANAGEMENT
186
184
VALID COMMANDS/RESPONSES

FIG. 6